Naval Research Laboratory

Washington, DC 20375-5320

NRL/FR/5740--97-9866

A Fuzzy Logic Multisensor Association Algorithm: Theory and Simulation

JAMES F. SMITH III

Surface Electronic Warfare Systems Branch
Tactical Electronic Warfare Division

September 30, 1997

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 30, 1997	Continuing

4. TITLE AND SUBTITLE

A Fuzzy Logic Multisensor Association Algorithm: Theory and Simulation

5. FUNDING NUMBERS

PE - 63270N

6. AUTHOR(S)

James F. Smith III

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/FR/5740--97-9866

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

A recursive multisensor association algorithm has been developed based on fuzzy logic. It simultaneously determines fuzzy grades of membership and fuzzy cluster centers. It is capable of associating data from various sensor types and performing without operator intervention. It associates data from the same target for multiple sensor types. The algorithm also provides an estimate of the number of targets present, reduced noise estimates of the quantities being measured, and a measure of confidence to assign to the data association. The fuzzy logic formalism used offers the opportunity to incorporate additional information or heuristic rules easily. A comparison of the algorithm to a more conventional Bayesian association algorithm is provided. Also, procedures for defuzzification, i.e., mapping fuzzy results to hard results are discussed as well as the method of determining target validity. Various simulated real-time data sets are analyzed and provide a basis for comparison of the fuzzy and Bayesian association algorithms.

14. SUBJECT TERMS

Fuzzy clustering	Recursive	Superclustering	Radar
Fuzzy sets	Multisensor	Outlier suppression	
Fuzzy logic	Association	ESM	

15. NUMBER OF PAGES

24

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18
298-102

CONTENTS

A FUZZY LOGIC MULTISENSOR ASSOCIATION ALGORITHM: THEORY AND SIMULATION

1. INTRODUCTION

The problem considered in this report is how to associate Electronic Support Measures (ESM) signals with one or more of m possible radar tracks. An algorithm based on fuzzy set theory has been developed to solve this problem. It considers the complexities offered by having multiple radar tracks and unequal numbers of measurements. It is capable of making its own estimate from ESM data of bearing and, as such, provides additional measures of association unlike the Trunk-Wilson (TW) Bayesian theory [1] with which it is compared. It can estimate the number of targets present in the data and use fuzzy set theoretic techniques to suppress outliers. The fuzzy grades of membership provide opportunities for incorporation of heuristic rule sets and extension to probability theory. The fuzzy cluster centers represent reduced noise estimates of the measured quantities. Finally, in comparison to an existing Bayesian algorithm, the fuzzy association algorithm exhibits superior performance.

In Section 2, the concepts of fuzzy set theory, hard and fuzzy clustering, defuzzification, and superclustering are introduced. Section 3 introduces the TW algorithm, an established algorithm with which the fuzzy algorithm will be compared. Section 4 discusses how, in a recursive algorithm, the process of estimating the number of targets present in the data through superclustering can be improved by introducing a priori information. Section 5 discusses fuzzy clustering results for simulated data and examines the performance of the fuzzy clustering and superclustering algorithms. Section 6 demonstrates the ability of the fuzzy association algorithm to deal with both noisy ESM and noisy radar. Section 7 discusses research and development related to the algorithm that will be carried out in the near future. Finally, Section 8 provides conclusions.

2. FUZZY SETS, CLUSTERING, DEFUZZIFICATION, AND SUPERCLUSTERING

The development of the fuzzy association algorithm requires the concepts of the fuzzy set, clustering, fuzzy clustering, defuzzification, and superclustering. These concepts are developed in the following subsections.

2.1 Fuzzy Set Theory

This section provides a basic introduction to the ideas of fuzzy set theory, which allows an object to have partial membership in more than one set. It does this through the introduction of a function known as the membership function, which maps from the complete set of objects X into a set known as membership space. More formally, the definition of a fuzzy set [2] is the following:

Manuscript approved September 4, 1997.

If **X** is a collection of objects denoted generically by x, then a fuzzy set **A** in **X** is a set of ordered pairs:

$$A = \{(x, \mu_A(x)) \mid x \in X\}.$$

$\mu_A(x)$ is called the membership function or grade of membership (also degree of compatibility or degree of truth) of x in A, which maps X to the membership space M. (When M contains only the two points 0 and 1, A is nonfuzzy, and $\mu_A(x)$ is identical to the characteristic function of a nonfuzzy set.) The range of the membership function is a subset of the non-negative real numbers whose supremum is finite. Elements with a zero degree of membership are normally not listed.

2.2 Fuzzy Clustering

Fundamental to the development of the fuzzy association algorithm is the concept of clustering. Clustering is an operation that allows data to be grouped into classes defined by a similarity measure. By definition [3], given K objects, the algorithm forms N clusters such that, with respect to the similarity measure, the members of each cluster have a greater similarity to each other than to the members of any other cluster.

The kind of clustering used for association is known as fuzzy clustering. A batch version of the algorithm appears in Ref. 4. There are many kinds of clustering. If each point is assigned 100% to a particular cluster, the algorithm is referred to as a hard-clustering algorithm. Fuzzy clustering differs from hard clustering in that each data point can have partial assignment in each cluster. The grade of membership previously defined under the concept of a fuzzy set gives the degree of membership of each point in each cluster.

In the development of the fuzzy association algorithm, clustering will play a significant role. The grades of membership will be established by minimizing a functional. This functional can be found in many places in the literature of fuzzy sets and fuzzy clustering [5,6]. It is defined below after some preliminary notation is established.

Let X be any finite set; V_{cn} is the set of real $c \times n$ matrices; c is an integer with $2 \le c \le n$, and n is the number of data points. The fuzzy c-partition space for X is the set

$$M_{fc} = \left\{ U \in V_{cn} \middle| u_{ik} \in [0,1] \forall i,k; \sum_{i=1}^{c} u_{ik} = 1 \forall k; 0 < \sum_{k=1}^{n} u_{ik} < n \forall i \right\}.$$

Row i of a matrix $U \in M_{fc}$ exhibits (values of) the ith membership function (or ith fuzzy subset) U_i in the fuzzy c-partition U of X. Stated less formally, u_{ij} is the grade of membership of data point j in fuzzy cluster i.

Definition: Let $J_m : M_{fc} \times R^{cp} \to R^+$,

$$J_m(U,v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m (d_{ik})^2,$$

where R^{cp} is the collection of possible p-dimensional vectors with real elements taken c at a time and R^+ is the real interval $[0,\infty)$;

$$U \in M_{fc}$$

is a fuzzy c-partition of X;

$$v = (v_1, v_2, \ldots, v_c) \in R^{cp} \qquad \text{with } v_i \in R^p$$

is the cluster center or prototype of u_i, $1 \le i \le c$;

$$(d_{ik})^2 = \|x_k - v_i\|^2 \qquad \text{and } \|\,.\,\|$$

is any inner product induced norm on R^p, and weighting exponent $m \in [1,\infty)$. Since each term of J_m is proportional to $(d_{ik})^2$, J_m is a square-error clustering criteria. The solution of the fuzzy clustering problem consists of minimizing J_m as a function of U and v subject to the constraints imposed in the definition of M_{fc}. Stated more formally, solutions of

$$\min_{M_{fc} \times R^{cp}} \{J_m(U,v)\}$$

are least-square error stationary points of J_m.

The value of m used for the simulations in Sections 5 and 6 is $m = 2$. The significance of the m parameter is more fully discussed in Ref. 6.

The goal of the fuzzy clustering algorithm is to determine fuzzy cluster centers v_i that represent the average value of the quantities in the fuzzy clusters, and the grade of membership of the kth data point in the ith fuzzy cluster for all data points-k and clusters-i. The algorithm determines these quantities by minimizing a least-square cost function where each term is weighted by a power of the grade of membership. Each term of the cost function simultaneously measures the distance of the data point from a cluster center and is weighted by the point's membership in that cluster. The minimization is conducted subject to the constraints that the sum of the grades of membership over clusters for a particular data point must equal unity, and for each cluster, the sum of grades of membership over data points must be bound between one and the maximum number of data points.

The fuzzy clustering algorithm requires input data to be clustered, the number of anticipated clusters, and an estimate of the grades of membership or the fuzzy cluster centers. The output will consists of:

- high quality estimates of the grades of membership, these quantities providing a measure of confidence of how well the data are clustered and a means of making an optimal data-point cluster assignment. This is especially useful if a data point falls on the boundary between clusters, and

- the fuzzy cluster centers, which will represent reduced noise values of the measured quantities. Cluster centers are also useful for conveying the position of the cluster.

For many applications, it is necessary to extract from the clustering algorithm a nonfuzzy, i.e., crisp statement of the assignment of each point. The process of taking fuzzy results and extracting definite, i.e., crisp data point-cluster assignment is known as *defuzzification.*

The current approach to defuzzification consists of making a definite data point assignment to that cluster for which the data point has the largest grade of membership. If it should occur that a data point has equal grades of membership for more than one cluster, the point, in the simplest form of defuzzification, is assigned to the first cluster that is encountered. A potentially better approach is discussed in Section 4.

2.3 Determination of the Fuzzy Grades of Membership and Fuzzy Cluster Centers

The cost function J_m is minimized over $M_{fc} \times R^{cp}$ by using Lagrange multipliers and taking derivatives. This gives rise to a coupled iterative system of equations. The fuzzy partition matrix and fuzzy cluster centers are iteratively changed until the norm of the change in the fuzzy partition matrix is less than a preset value. When this convergence criterion is applied to the values of the fuzzy partition matrix, the coupled system and the convergence criterion are referred to as the Picard algorithm. The Picard algorithm is guaranteed to converge to a local minimum; this particular type of fuzzy clustering is referred to as a c-means algorithm [5].

The system of equations resulting from the minimization represents a coupling between the fuzzy cluster centers and the fuzzy partition matrix. An initial estimate of either quantity is all that is required to initialize the iteration process. Thus, an initial estimate of the fuzzy cluster center by one class of sensors can be used to cluster data measured by another sensor system. The algorithm can be made recursive by using previous estimates of cluster centers and/or cluster center estimates derived from other sensors.

Another procedure for initializing the iterative process is to initially estimate the fuzzy partition matrix, i.e., the grades of membership, using some other clustering algorithm. Since the Picard algorithm is guaranteed to converge to a local minimum, if the fuzzy clustering algorithm is initialized using a good but not perfect clustering algorithm, it can frequently improve clustering results because of its ability to deal with ambiguous data-point cluster assignments.

2.4 Superclustering

Clustering algorithms, including the fuzzy clustering algorithm, generally require a specification of the final number of clusters. If the data being clustered represent ships, aircraft, missiles, etc., this implies a priori knowledge of the number of targets. Obviously, in general, the number of targets will not be known before processing. So it is desirable to develop a technique for determining from the data the appropriate number of clusters, i.e., the number of targets. Such a technique, known as *superclustering*, has been developed that provides a solution to this problem. The superclustering techniques developed here are related to and represent an extension of techniques in fuzzy cluster validity theory [5].

The method of superclustering is described as follows: given an upper bound on the number of clusters, this bound is supplied to the fuzzy clustering algorithm. This algorithm produces this number of clusters for the data with associated grades of membership for each data point in each cluster. The fuzzy clustering algorithm also provides the coordinates of the fuzzy cluster centers. Intuitively, clusters should be separated, nonoverlapping, and not extremely close to each other with respect to some

measure. It then becomes essential to define a measure of "closeness" and provide a criterion for what "too close" means.

An obvious candidate for a measure of closeness of two clusters is the separation of the cluster centers. The cluster centers do not tell the whole story. The data points may be distributed close to the cluster center or they may be a significant, absolute distance from it. Also, when dealing with fuzzy clustering (before defuzzification), the points generally do not belong 100% to any cluster. In an effort to provide a unitless measure of closeness and incorporate the concept of vagueness inherent in fuzzy algorithms, the distance between fuzzy cluster centers should be normalized by some function of the grades of membership. Incorporation of the grades of membership, i.e., superclustering before defuzzification, has the advantage of potentially better cluster assignments for points that fall on the boundary between clusters.

One such normalized measure of cluster center separation is the *c-matrix*, defined below. Let *v(i)* and *v(j)* be the position vectors for the fuzzy cluster centers for cluster *i* and cluster *j*, respectively, and *N* the number of data points. Then the *i*th – *j*th element of the *c-matrix* is

$$c(i,j) = \| v(i) - v(j) \| / \max(std(i), std(j)), \tag{1}$$

where

$$std(k) = \sqrt{\sum_{i=1}^{N} u(i,k)^m * (x(i) - mean(k))^2 \Big/ \sum_{i=1}^{N} u(i,k)^m}, \tag{2}$$

and

$$mean(k) = \left(\sum_{i=1}^{N} u(i,k)^m * x(i) \right) \Big/ \sum_{i=1}^{N} u(i,k)^m. \tag{3}$$

Equations (2) and (3) define the fuzzy standard deviation and the fuzzy mean, respectively.

The *c-matrix* capitalizes on the intuitive idea that cluster centers should be separated by a certain number of fuzzy standard deviations. If cluster centers are closer than this, they probably correspond to the same cluster. If it is determined that two or more clusters should be merged into a single cluster, the resulting grouping will be referred to as a *supercluster*. A criterion must be established to determine when supercluster formation is warranted. A simple criterion consists of defining a threshold τ such that if $c(i,j) < \tau$, then clusters *i* and *j* are merged into a supercluster. A method of selecting the value of τ is discussed below.

A simple criterion for selecting the value of τ would be to first consider the elements of each cluster as points randomly distributed around some mean value. If the data have a Gaussian distribution, then 98% of the points are within three standard deviations of the mean. So a value of $\tau = 3$ is selected.

After *c-matrix* formation and establishing the threshold, the next step involved in superclustering is determining exactly how to form superclusters, i.e., when there is more than one choice based on what has been developed up to now, what is the best supercluster formation scheme. Accordingly, four different procedures for supercluster formation have been examined [4].

Superclustering is conducted as follows: first fuzzy clustering is carried out followed by *c-matrix* formation. All the fuzzy cluster centers that are within threshold of fuzzy cluster center one fall into

the first supercluster. Those that are within threshold of fuzzy cluster center two fall into the second supercluster and so on. This allows determination of the number of superclusters—hence, targets. New fuzzy cluster centers and grades of membership are calculated corresponding to the determined number of superclusters using the fuzzy clustering algorithm described above. In Ref. 4, this algorithm is compared to several other more sophisticated superclustering schemes and found to be acceptable except under conditions of very poor data.

3. THE TRUNK-WILSON (TW) ALGORITHM

The TW algorithm assumes there are K ESM tracks, each specified by a different number of ESM measurements. These ESM measurements will be associated with either no radar track or one of m radar tracks, each radar track having a different number of measurements. The association of ESM tracks with radar tracks using the multiple hypothesis testing technique is as follows:

H_0: ESM measurements are associated with no radar track;
H_1: ESM measurements are associated with the first radar track;

.
.
.

H_j: ESM measurements are associated with the jth radar track;

.
.
.

H_m: ESM measurements are associated with the mth radar track.

The Bayesian procedure, which minimizes the probability of error, is to select the hypothesis having the largest a posteriori probability. It is assumed that the ESM measurement errors are independent and Gaussian distributed with zero mean and constant variance σ^2. If the a priori probabilities are equal, the minimum error decision rule selects the target j based on the statistical distance d_j for which d_j is minimized and given by

$$d_j = \sum_{i=1}^{K_j} \left[\theta_e(t_i) - \theta_j(t_i)\right]^2 / \sigma^2 \qquad j = 1, \ldots, m, \tag{4}$$

where $\{\theta_e(t_i),\ 1 \le i \le K_j\}$ is a set of K_j ESM bearing measurements, and $\theta_j(t_i)$ is the true bearing to target j at the time of the ith ESM measurement. The minimized statistical distance d_{jmin} has a chi-square density with K_j degrees of freedom. Consequently, the desired a posteriori probability P_j is given by

$$P_j = \int_{d_j}^{\infty} \chi^2(t)\,dt, \tag{5}$$

where $\chi^2(t)$ is the chi-square density function. The TW algorithm selects the association that has the largest probability, P_{max} where

$$P_{max} = \max\left\{P_j \mid j = 1, \ldots m\right\}.$$

The TW algorithm also uses the next largest probability P_{next} and four decision theoretic quantities obtained by Trunk and Wilson. Three of the quantities are the high (T_H), middle (T_M), and low (T_L) probability thresholds. The fourth quantity is the probability margin (R). The corresponding decision rules are:

(1) firm correlation, $P_{\max} \geq T_H$ and, $P_{\max} \geq P_{next} + R$, the ESM signal goes with the radar track having largest P_j (i.e., $P_{\max}$),

(2) tentative correlation, $T_H > P_{\max} \geq T_M$ and $P_{\max} \geq P_{next} + R$, ESM signal probably goes with radar track having largest P_j (i.e., $P_{\max}$),

(3) tentative correlation with some track, $P_{\max} \geq T_M$ but $P_{\max} < P_{next} + R$, ESM signal probably goes with some radar tracks (but the algorithm cannot determine which),

(4) tentatively uncorrelated, $T_M > P_{\max} \geq T_L$, ESM signal probably does not go with any radar track, and

(5) firmly uncorrelated, $T_L \geq P_{\max}$, ESM signal does not go with any radar track.

The threshold T_H is set equal to P_{FA} defined as the probability of falsely associating a radar track with an ESM signal when the ESM signal does not belong with the radar track. The threshold T_H is a function of the azimuthal difference, denoted by μ, between the true (ESM) position and the radar track under consideration. The threshold T_H was determined for two values $\mu = 1.0\sigma$ and 1.5σ by Trunk and Wilson using simulation techniques. Also, the threshold T_M was determined using the simulation previously used to determine T_H for the same two values of μ, 1.0σ and 1.5σ. The threshold T_L is defined as a rejection rate of P_R. In this report, $\mu = 1.5\sigma$ is used, and the threshold T_L is set equal to 0.001.

The probability margin R is determined by specifying a probability of an association error P_e according to the following equation

$$P_e = P_{\mathrm{R}} \left\{ P_{\max} \geq P_{\mathrm{next}} + R \right\},$$

where $P_{\max}$ corresponds to an incorrect association, and P_{next} corresponds to the correct association. The probability margin R is a function of P_e and the separation μ of the radar tracks. The values of R were determined by Trunk and Wilson for $P_e = 0.01$ by simulation techniques.

4. PROBABILITY AUGMENTED SUPERCLUSTERING

In many cases, superclustering alone can determine the number of targets exactly or within one target if there are enough data [4]. In a recursive algorithm, data are arriving from one moment to the next. The amount of data required to make a correct decision as to the number of targets can be reduced by adding in a priori information, e.g., noise statistics. By calculating the probability that the data are associated with each cluster center, cluster centers with a probability of association much less than the maximum can be neglected. The rejection threshold that has been found to be useful is 20% of the maximum probability. This procedure is referred to as *probability augmented superclustering* and has been found to be quite effective (Section 5).

5. APPLICATION OF THE FUZZY ASSOCIATION ALGORITHM TO SIMULATED DATA AND COMPARISON TO THE TW ALGORITHM

5.1 Bearing Estimation and Superclustering

This section demonstrates the fuzzy algorithm's ability to estimate parameters and determine the number of targets present in the simulated ESM data. The ESM bearing data were simulated for a single target moving with constant bearing of 0° with zero mean, 1° standard deviation Gaussian noise added.

Figure 1 is a plot of the fuzzy estimates of bearing in degrees versus the number of data points for a simulated target having constant bearing of 0°. The algorithm is initialized using the radar estimates given in Table 1 for Example 1. The algorithm is not very sensitive to initialization, thus Figs. 1 and 2 are similar if the radar estimates in Examples 1 through 4 are used. The absolute fuzzy bearing estimate is always less than 0.12°; for more than four data points, it is always less than 0.07°. The fuzzy association algorithm is initialized using radar estimates. The fuzzy algorithm converges rapidly and, as such, the bearing estimate curves look similar, independent of initialization. Due to the similarity of the figures, only one will be displayed to save space.

In this example, the fuzzy algorithm estimates the value of bearing with an error of no more than 0.12°, even with only two data points. The error falls off rapidly with additional data points. As the number of data points approaches 50, the error approaches zero, as expected.

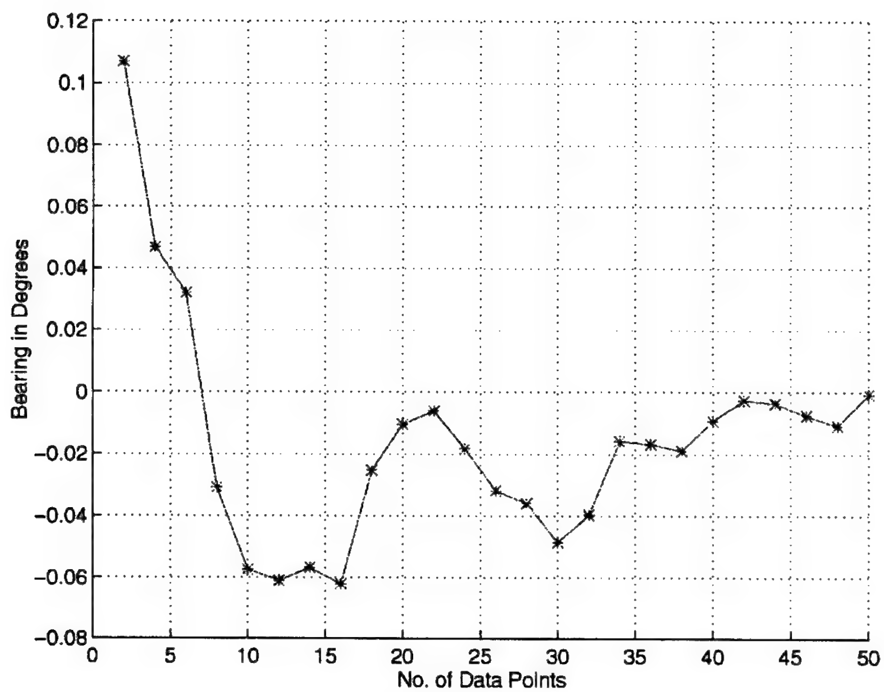
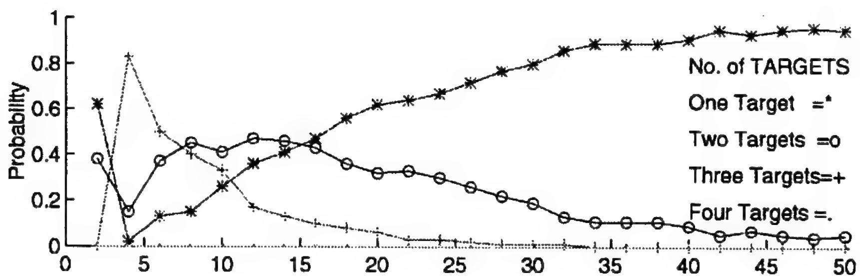


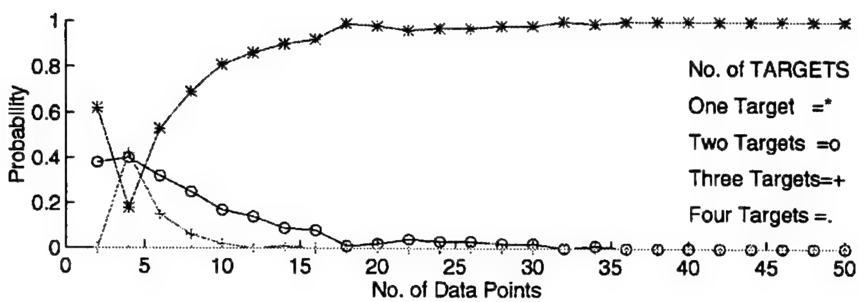
Fig. 1 – Fuzzy bearing estimate in degrees vs number of data points. Truth = 0° and one target.

Table 1 – Radar Measurements

Radar Simulation Cases			
Example	*a*	*b*	*c*
1	0°	1°	–1°
2	2°	1°	–1°
3	0°	2°	–2°
4	4°	2°	–2°



(a)



(b)

Fig. 2 – The probability of (a) superclustering determining there are one, two, three, or four targets and (b) the probability of probability-augmented superclustering determining there are one, two, three, or four targets

Figure 2(a) is the probability versus the number of data points that superclustering determines the following number of targets: one target (*), two targets (o), three targets (+), or four targets (.). In this example, there must be more than 15 data points for superclustering to determine the exact number of targets more than 50% of the time. After 30 data points, superclustering is correct more than 80% of the time. The probability of superclustering making a two-target error remains less than 10% by the time 16 data points have been accumulated. The probability of the estimate being off by one target falls below 30% by 25 data points. As for the bearing estimates, the plots for this case are similar, and as such, only one will be displayed.

The observation that superclustering can be off by one or two targets for small numbers of data points suggests that a few outliers are establishing their own cluster centers. When additional data points are added, the effect of these additional cluster centers is small resulting in those data points being superclustered into a larger supercluster, thereby improving the estimate of the number of targets. The fact that, for small number of data points, some outliers establish their own superclusters

suggests that an outlier suppression mechanism can be usefully employed. It is this motivation that gives rise to the probability-augmented superclustering procedure that is used in Fig. 2(b).

Figure 2(b) is a plot of probability-augmented superclustering determining that there are one target (*), two targets (o), three targets (+), or four targets (.). Once again, the vertical axis is the ensemble probability, and the horizontal axis is the number of data points used. By the time six data points are accumulated, there is a 50% probability that the number of targets will be predicted correctly and only a 10% chance that the number of targets estimated will be off by two. When 10 data points are collected, the probability of correctly estimating the number of targets exceeds 80%. With more than 16 data points, the probability of correctly estimating the number of targets exceeds 90% and remains nearly 100% for 20 or more data points.

The use of the probability augmentation test to suppress outlier superclusters greatly improves the quality of the estimates of the number of targets present. In a comparison between Figs. 2(a) and 2(b) for 4 to 10 data points, the probability-augmented test increases the probability of correctly estimating the number of targets by 10% to 30%. For 16 data points, the improvement is about 40%.

5.2 Comparison of the Fuzzy and TW Association Algorithms

In this section, four simulations are considered (Examples 1 to 4). In each case, the target has a constant bearing of 0°, and there are three constant-bearing radar estimates (a, b, and c). Table 1 shows the radar estimates, which provide the distinction among the four examples. In each simulation, 2 data points are added each time until total of 50 data points are accumulated. The ESM data used are described in Section 5.1. For each example, the simulation has been run 1000 times, and the results averaged, i.e., an ensemble with 1000 elements is used.

Figures 3 through 10 summarize the results of simulations that compare the ability of the fuzzy association algorithm using probability-augmented superclustering and the TW algorithm. There are four simulation examples corresponding to the four different radar-bearing tracks. The radar measurements are noiseless. The data were generated as in Figs. 1 and 2. As above, the data represent a target moving with constant bearing of 0° with zero mean unit variance Gaussian noise added. The four sets of radar measurements each consist of three bearing measurements that are constant in time.

The algorithms determine one of the following about the ESM data, as explained in Section 3. The ESM data are

- firmly correlated (FCT) with one of the radar tracks,
- tentatively correlated with one of the radar tracks (TCT),
- tentatively correlated with an unknown radar track (TUT),
- tentatively uncorrelated with any of the radar tracks (TNT), or
- firmly uncorrelated with the radar tracks (FNT).

The fuzzy association algorithm compares its estimate of bearing to the radar measurements and uses the fuzzy standard deviation for σ. So when Eq. (4) is used in the fuzzy association algorithm,

$\theta_e(t_i)$ = the fuzzy bearing estimate at the time t_i and
σ = the fuzzy standard deviation.

The substitution of the fuzzy bearing estimate and the fuzzy standard deviation into Eq. (4) is what distinguishes the fuzzy association algorithm from the TW association algorithm. All definitions of correlation, i.e., FCT, TCT, TUT, TNT, and FNT remain the same.

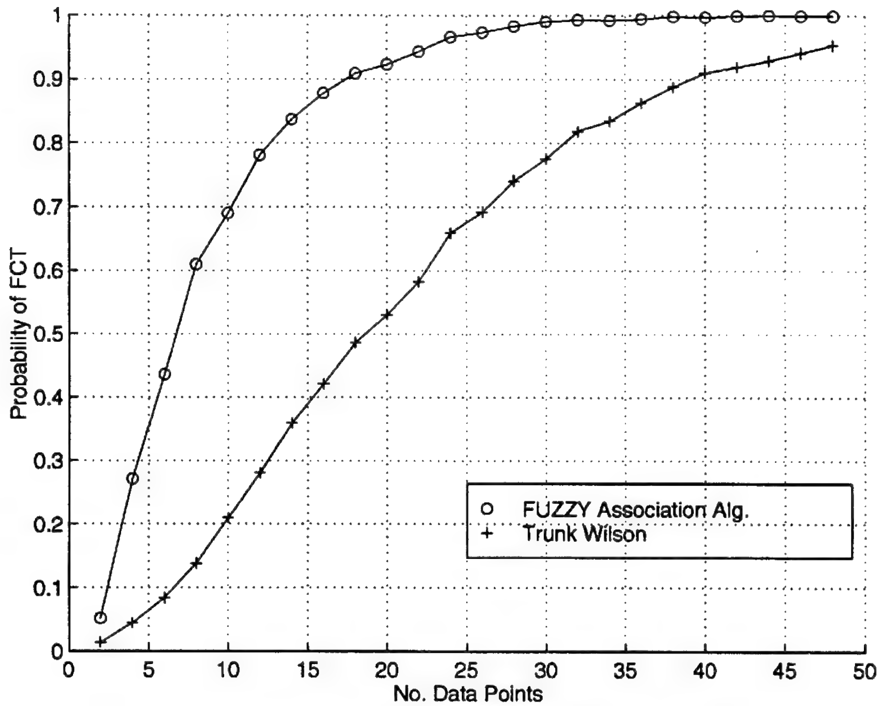


Fig. 3 – The fuzzy association and the TW algorithms probability of FCT for the radar example 0°, 1°, –1°. Truth is 0°.

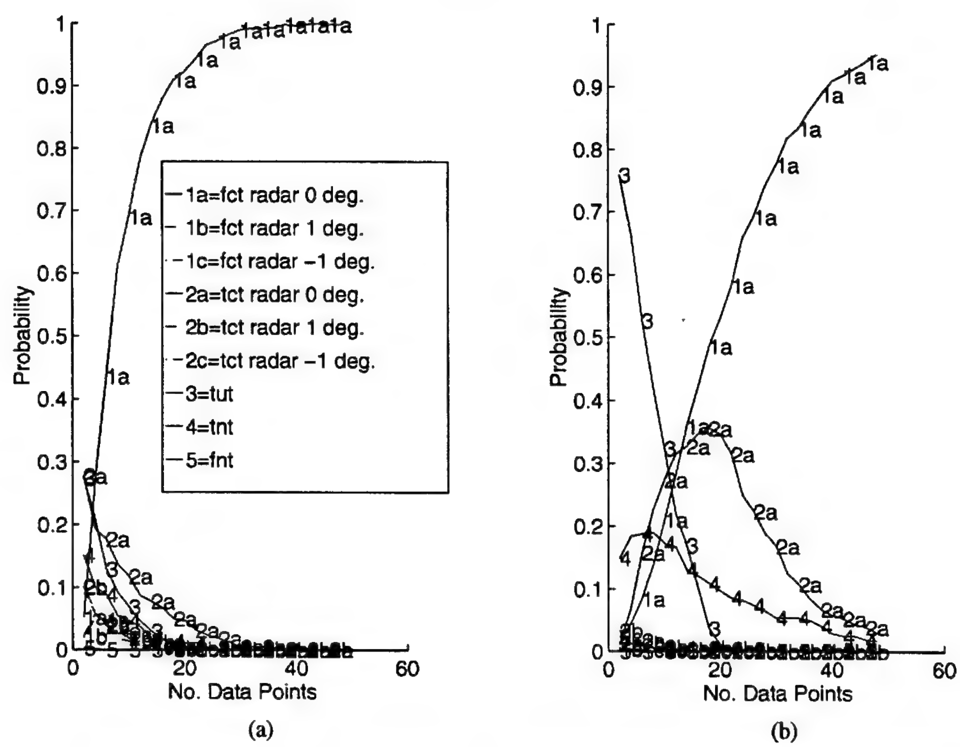


Fig. 4 – Radar example 0°, 1°, –1°: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

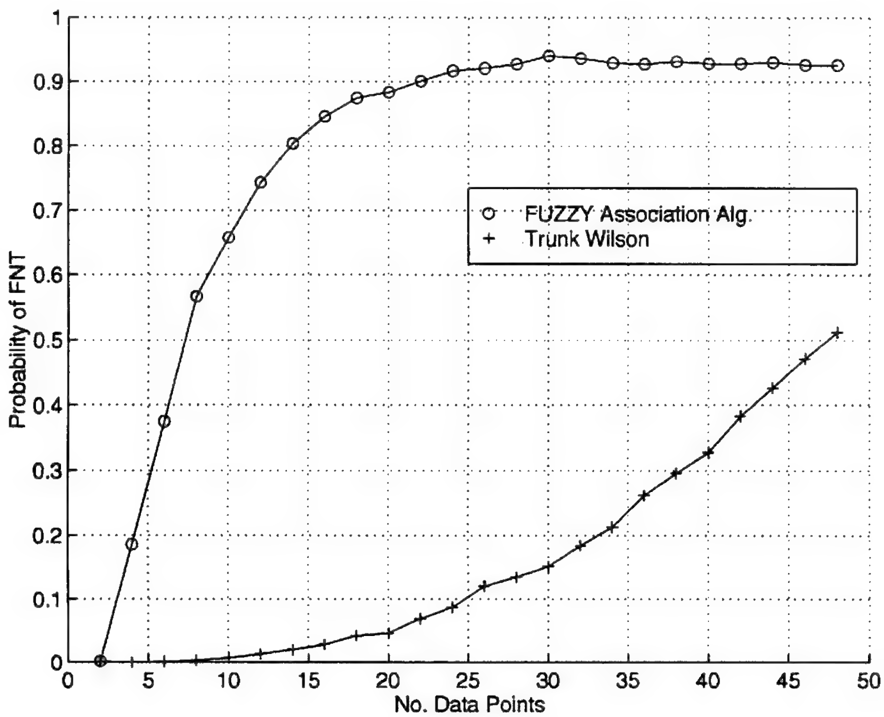


Fig. 5 – The fuzzy association and the TW algorithms probability of FNT for the radar example 2°, 1°, –1°. Truth is 0°.

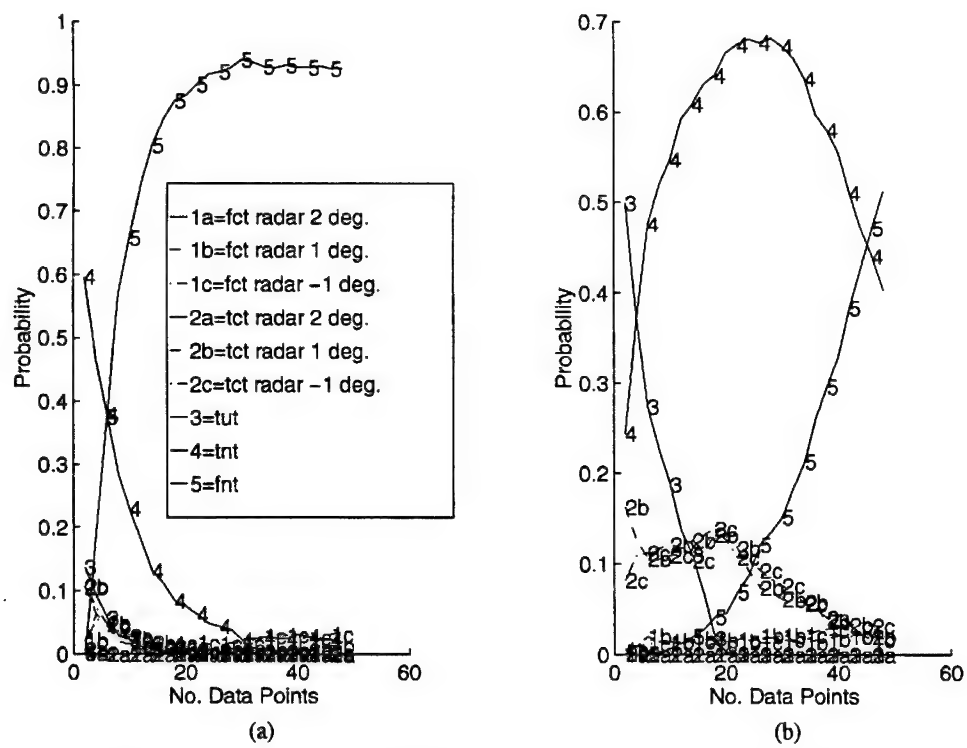


Fig. 6 – Radar example 2°, 1°, –1°: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

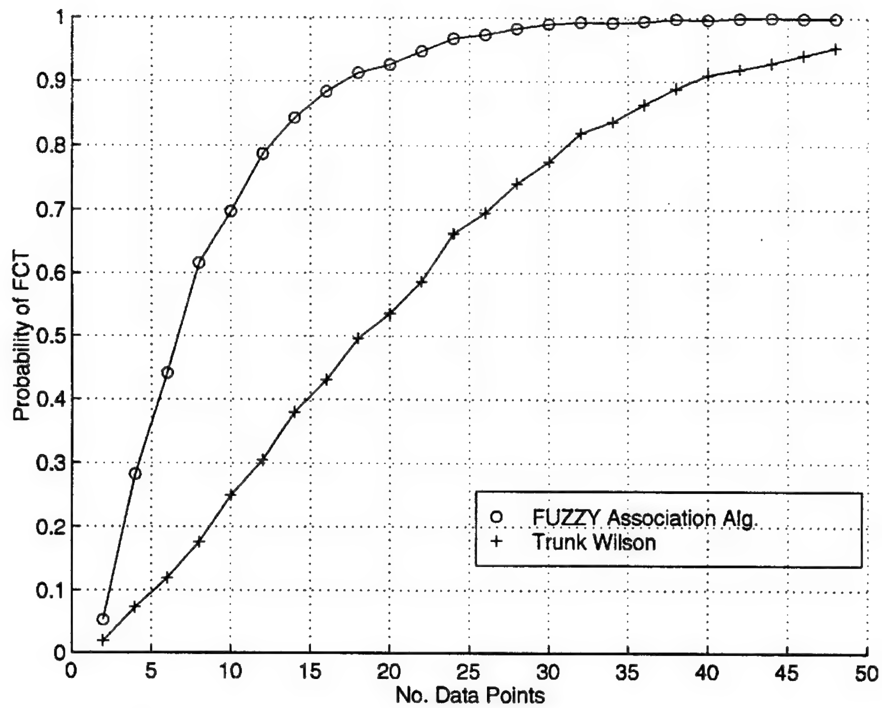


Fig. 7 – The fuzzy association and the TW algorithms probability of FCT for the radar example 0°, 2°, –2°. Truth is 0°.

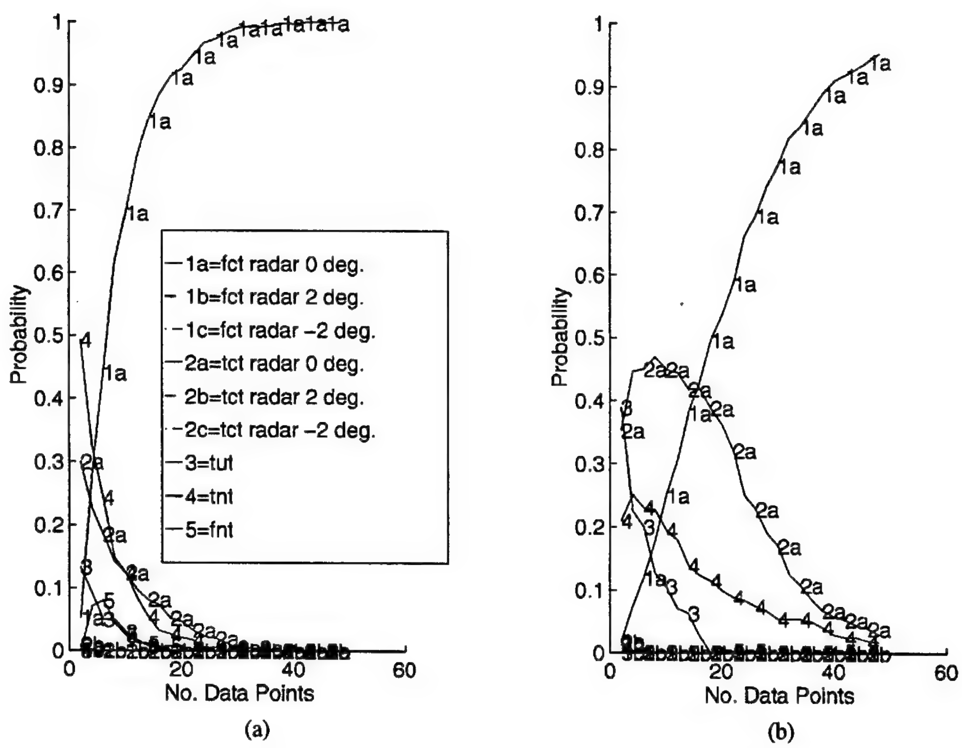


Fig. 8 – Radar example 0°, 2°, –2°: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

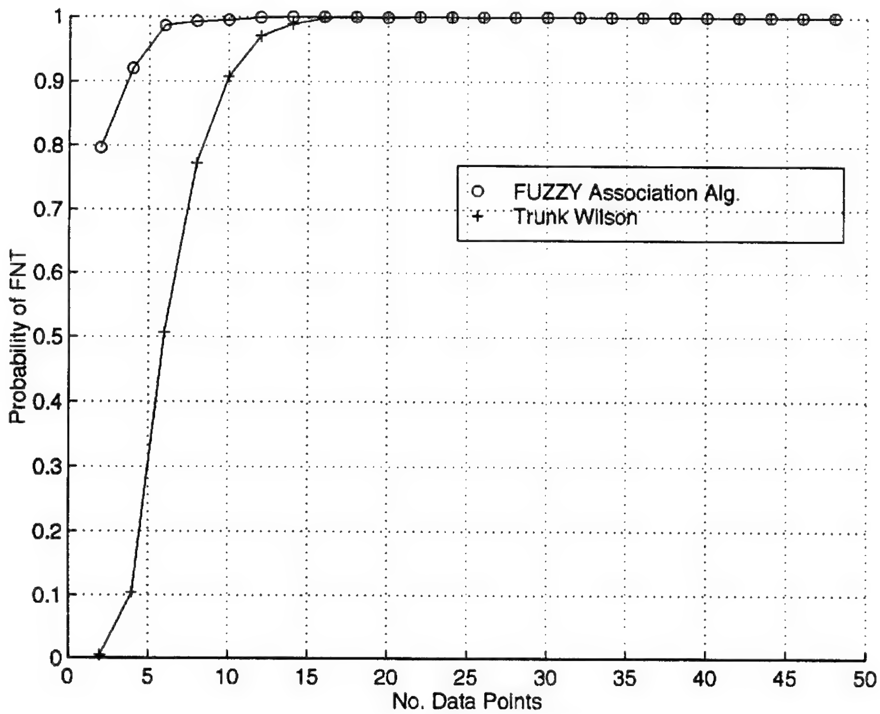


Fig. 9 – The fuzzy association and the TW algorithms probability of FNT for the radar example 4°, 2°, –2°. Truth is 0°.

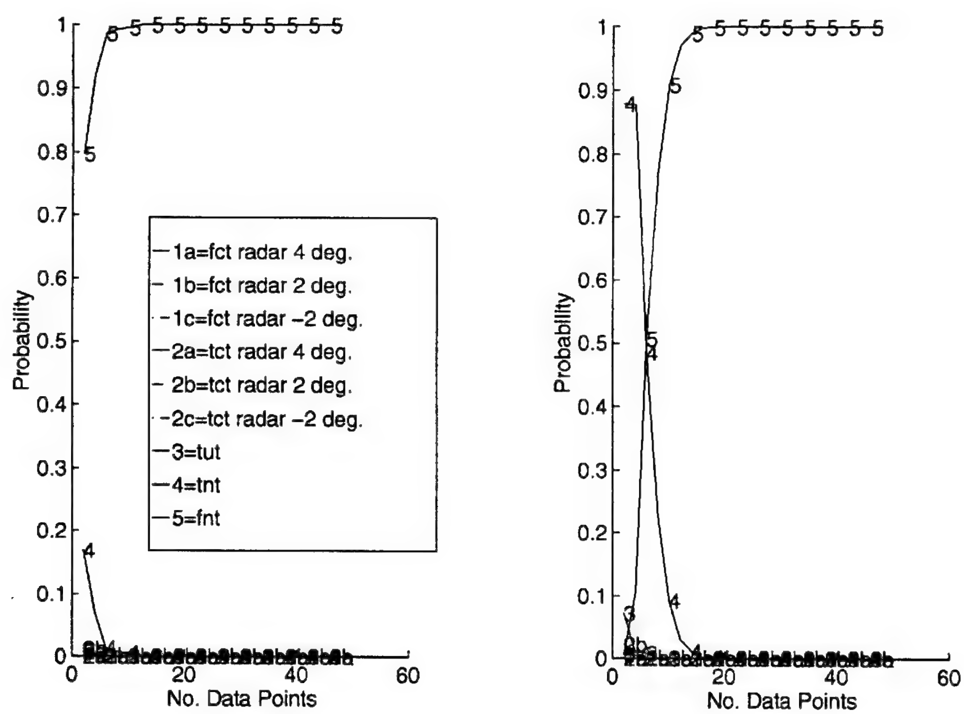


Fig. 10 – Radar example 4°, 2°, –2°: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

The figures summarize the probability of occurrence of each hypothesis class. Since in each example there are three radar tracks for potential association with this ESM data, it follows there are 10 potential probability curves to plot. Figures 4, 6, 8, and 10 exhibit all 10 curves for Examples 1 through 4. The fuzzy association algorithm results are always plotted in the left figure (a) and the TW results, in the right figure (b).

Figures 3, 5, 7, and 9 compare the relevant curves from the fuzzy association algorithm and the TW algorithm on the same axes. The fuzzy association algorithm results are given by the curve marked with *o*'s, and the TW results are indicated by the curve marked with *+'s*. The vertical axis indicates probability of FCT (Figs. 3 and 7) or FNT (Figs. 5 and 9) and the horizontal axis, the number of data points necessary to establish the associated level of probability.

In Figs. 4, 6, 8, and 10, the results for all hypothesis classes for the fuzzy algorithm (Figs. 4(a), 6(a), 8(a), 10(a)) and the TW algorithm (Figs. 4(b), 6(b), 8(b), 10(b)) are displayed using the following labeling scheme. In both (a) and (b), there are 10 curves. The data represent the probability that each of the hypothesis classes occur versus the number of data points. The legend indicates the meaning of each of the symbols. A common system of marking has been employed for figure types (a) and (b). Curves indicating probability of FCT and TCT are marked with 1's and 2's, respectively. The curves that also have an *a, b,* or *c* label, correspond to the appropriate radar track indicated in Table 1. The remaining three hypothesis classes—TUT, TNT, and FNT—are marked with a 3, 4, and 5, respectively. The curves in Figs. 4(a), 6(a), 8(a), and 10(a) marked with a "1(a)" correspond to the curves marked with a "+" in Figs. 3, 5, 7, and 9. Curves marked with a "1(a)" label in the even "b" figures correspond to curves marked with an "o" in the odd numbered figures.

Figure 3 presents results for the radar example $\mu = 0°, 1°, -1°$. Since the radar results contain truth, i.e., a target moving with constant bearing of 0°, a good association algorithm will establish that there is a firm correlation between the ESM data and the 0° bearing track. Figure 3 plots the probability that the association algorithms establish a firm association between ESM data and the radar measurements. The fuzzy association algorithm results are given by the curve marked with *o*'s, and the TW results are indicated by the curve marked with *+'s*. The vertical axis indicates probability of firm correlation, and the horizontal axis, the number of data points necessary to establish that level of probability.

The fuzzy association algorithm results are always superior to the TW algorithm. At 10 data points, the fuzzy algorithm has established a 70% probability of firm correlation, whereas the TW algorithm requires about 26 points to establish the same level of probability of FCT. An 80% probability of FCT is established by the fuzzy algorithm by the 13th data point, whereas TW requires about 32 points to reach the same level of success. The fuzzy algorithm reaches 90% probability of FCT at 18 data points and the TW algorithm at about the 39th point. So the fuzzy algorithm establishes high probabilities of firm correlation with one-third to one-half the data required by the TW algorithm. In this sense, the fuzzy algorithm is two to three times faster than the TW algorithm. Also, this is a difficult example for any association algorithm since there are two additional radar measurements within one noise standard deviation. The ability of the fuzzy algorithm to make high-quality decisions with fewer data points than the TW algorithm is significant since real data are frequently sparse and intermittent.

Figure 4 considers the same ESM and radar data as Fig. 3, with all hypothesis classes for the fuzzy algorithm (a) and the TW algorithm (b) plotted for completeness. In Fig. 4(a), the fuzzy association algorithm's probability of firm correlation with 0° radar measurement rises rapidly. The other hypothesis classes approach zero within only a small number of data points. After 12 data points, all

other hypothesis classes have probability of < 10%. In comparison, the probability of firm correlation between ESM and the 0° radar track has a much slower rise for the TW algorithm, with the other hypothesis classes maintaining high probabilities. Even at 20 data points, the TW algorithm gives a 20% to 30% probability that the ESM, and 0° radar track are only tentatively associated. With fewer than 20 data points, the TW algorithm gives significant probabilities of declaring that the data are TUT or TNT. In conclusion, the TW results are much more ambiguous than the fuzzy association results.

Figure 5 presents results for the radar example $\mu = 2°, 1°, -1°$. Since the radar results do not contain truth, i.e., a target moving with constant bearing of 0°, a good association algorithm will establish that the ESM data are firmly uncorrelated with the radar tracks. Figure 5 plots the probability the association algorithms establish that the ESM and radar data are firmly uncorrelated. By the 8th data point, the fuzzy algorithm has reached a 55% probability of FNT, whereas the TW algorithm never exceeds that probability. For this example, the fuzzy algorithm is six times faster than the TW algorithm, i.e., it reaches the TW algorithm's maximum probability of FNT with one-sixth of the data. The fuzzy algorithm has 80% and 90% probability of FNT by the 13th and 20th data points, respectively. Thus once again, the fuzzy algorithm makes a high-quality decision long before the TW algorithm.

Figure 6 considers the same ESM and radar data as Fig. 5. Figures 6(a) and 6(b) give all 10 hypothesis classes for the fuzzy algorithm (a) and the TW algorithm (b). In Fig. 6(a), the fuzzy association algorithm's probability of FNT rises rapidly. The other hypothesis classes approach zero within a small number of data points. After five data points, only the probability that the ESM and radar data are TNT or FNT are still significant. The probability that the data are tentatively not correlated has fallen below 10% by 12 data points, and the probability of firm correlation is above 80% by 14 data points. In comparison, in Fig. 6(b), the TW algorithm never has more than a 54% probability of declaring that the ESM and radar data are FNT. For the first 20 data points, there are significant probabilities of this algorithm declaring five different hypothesis classes to be correct. So once again, the conclusions of the TW algorithm are more ambiguous than the fuzzy association algorithm.

Figure 7 presents results for the radar example $\mu = 0°, 2°, -2°$. Since the radar results contain truth, i.e., a target moving with constant bearing of 0°, a good association algorithm will establish that there is a firm correlation between the ESM data and the 0° bearing track. The fuzzy association algorithm results are always superior to the TW algorithm. At 10 data points, the fuzzy algorithm has established a 70% probability of firm correlation, whereas the TW algorithm requires about 26 points to establish the same level of probability of FCT. An 80% probability of FCT is established by the fuzzy algorithm by the 13th data point, whereas the TW requires about 32 points to reach the same level of success. The fuzzy algorithm reaches 90% probability of FCT at 16 data points and the TW algorithm at about the 38th point. So the fuzzy algorithm establishes high probabilities of firm correlation with one-third to one-half the data required by the TW algorithm. In this sense, the fuzzy algorithm is two to three times faster than the TW algorithm.

Figure 8 considers the same data in Fig. 7, except as in Fig. 6, all hypothesis classes are displayed. In Fig. 8(a), the fuzzy association algorithm's probability of firm correlation with the 0° radar track rises rapidly. The other hypothesis classes approach zero within only a small number of data points. After 12 data points, all other hypothesis classes have probability of less than 10%. The 0° FCT TW curve exhibits a much slower rise with the other TW hypothesis classes maintaining correspondingly high values. Even at 20 data points, the TW algorithm gives about a 40% probability that the ESM and 0° radar track are only tentatively associated. With fewer than 20 data points, the TW algorithm gives significant probabilities of declaring that the data are TUT or TNT. So in conclusion, the TW results are more ambiguous than the fuzzy association results.

Figure 9 presents results for the radar example $\mu = 4°, 2°, -2°$. Since the radar results do not contain truth, i.e., a target moving with constant bearing of 0°, a good association algorithm will establish that the ESM data are firmly uncorrelated with the radar tracks, i.e., it is FNT. Figure 5 plots the probability the association algorithms establish that the ESM and radar data are firmly uncorrelated. By two data points, the fuzzy algorithm has established an 80% probability of FCT and a 90% probability of FNT by the fourth data point. The TW algorithm establishes an 80% probability of FNT by about the 8th data points and a 90% probability of FNT by 10th point. This is an extremely easy example for the algorithm to determine that the radar data are firmly uncorrelated with the ESM data since the radar measurements are two to four noise standard deviations from truth.

Figure 10 considers the same example as Fig. 9, except that all the hypothesis classes are displayed. In Fig. 10(a), the fuzzy association algorithm's probability of FNT rises rapidly. The other hypothesis classes approach zero within only a small number of data points. After two data points, only the probability that the ESM and radar data are firmly uncorrelated is significant. The probability of FNT is above 90% by four data points. In comparison, in Fig. 10(b), the TW algorithm reaches a 90% probability of FCT by 10 data points. Like the fuzzy association algorithm, only the probability of FNT remains significant for the TW algorithm after 10 data points. In this example, all radar estimates are two to four standard deviations from truth. This accounts for the ability of both algorithms to come to a correct conclusion with so few points. This example is easier than the others. Observe that even for an example that is very easy for the TW algorithm, the fuzzy algorithm comes to a correct conclusion much more rapidly.

6. ASSOCIATION OF NOISY ESM AND NOISY RADAR MEASUREMENTS

In the previous simulations, it was assumed that the radar measurements were noiseless. Following the procedure of Ref. 8, the TW algorithm can be used to associate noisy ESM and noisy radar measurements as follows. The radar measurements for radar track j at time t_i will have zero mean Gaussian noise added to them. The variance of the noise will be denoted as σ_{ij}^2 for the jth radar track and the ith time. If the variance in Eq. (4) is replaced by

$$\sigma^2 = \sigma_E^2 + \sigma_{ij}^2 ,$$

where $\sigma_E^2 =$ the variance of the ESM noise, then the statistic defined in Eq. (4) still has a chi-square density. Thus, it follows that all the thresholds should have the same value, whether or not the radar measurements are noisy. There will be some difference in multitarget performance because of the different dependence between the squared errors $\{d_j, j = 1,2,\ldots,m\}$ due to the radar variances.

Figures 11 and 12 are for the radar example 0°, 1°, −1° with $\sigma_{ij} = 0.1°$ for all times t_i and radar tracks j. They correspond to the perfect radar cases given in Figs. 3 and 4. The radar noise standard deviation is consistent with levels found in modern radar systems. Figures 3 and 11 are practically indistinguishable, as are Figs. 4 and 12. This implies that both algorithms are insensitive to the small amounts of noise found in modern radar systems.

Figures 13 and 14 correspond to radar example 2°, 1° −1°. Once again the radar noise has standard deviation of $\sigma_{ij} = 0.1°$ for all times t_i and radar tracks j. Figures 13 and 14 correspond to the noiseless cases provided in Figs. 5 and 6. Once again, both the fuzzy association and TW association algorithm show little sensitivity to the radar noise.

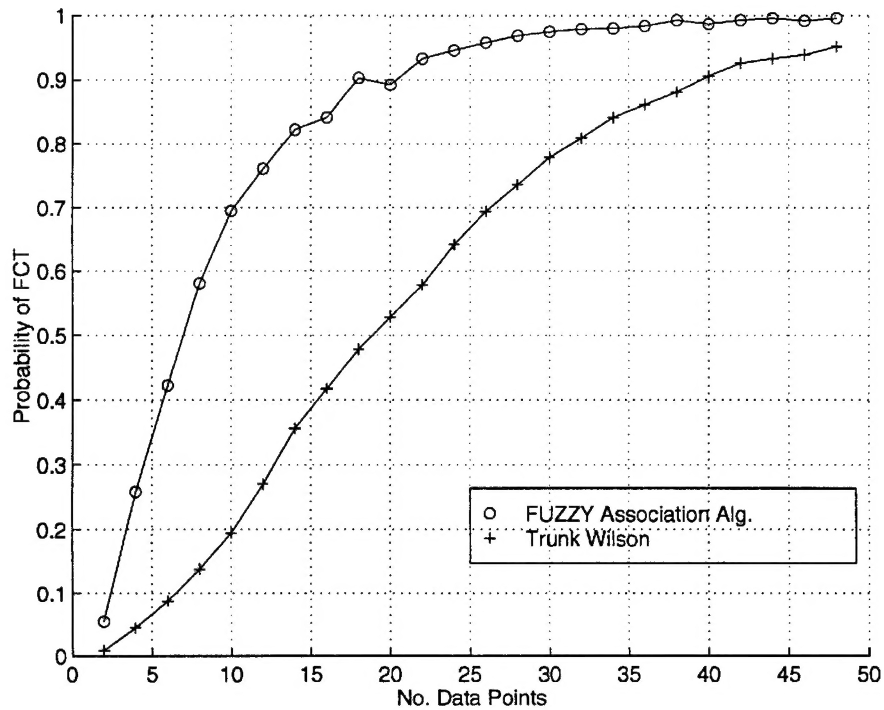


Fig. 11 – Radar example 0°, 1°, –1° with σ_{ij} = 0.1° for all times t_i and radar tracks j

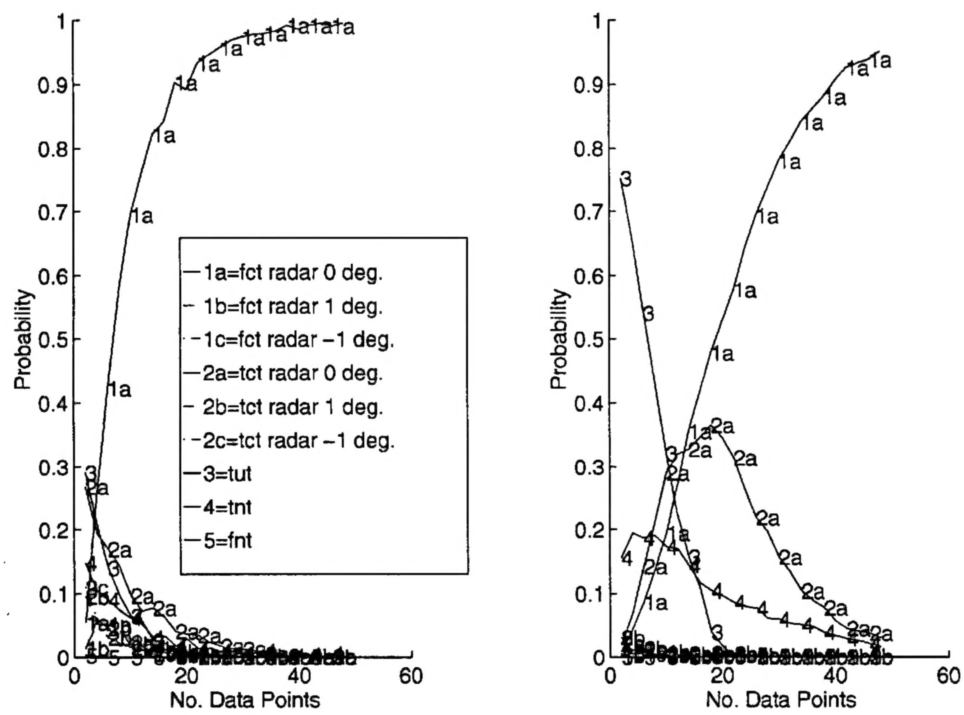


Fig. 12 – Radar example 0°, 1°, –1° with σ_{ij} = 0.1° for all times t_i and radar tracks j: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

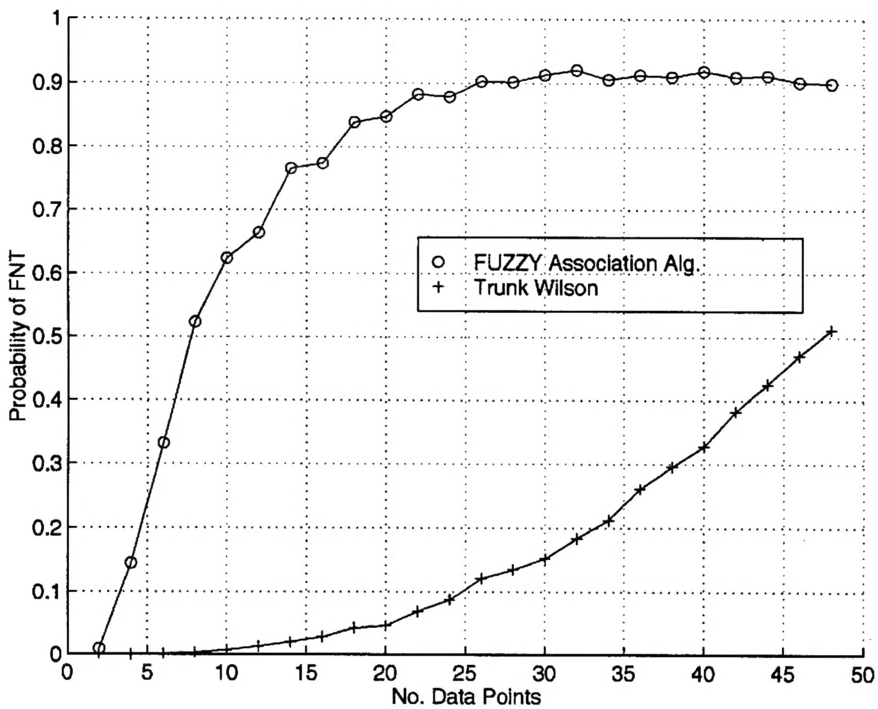


Fig. 13 – Radar example 2°, 1°, -1° with σ_{ij} = 0.1° for all times t_i and radar tracks j

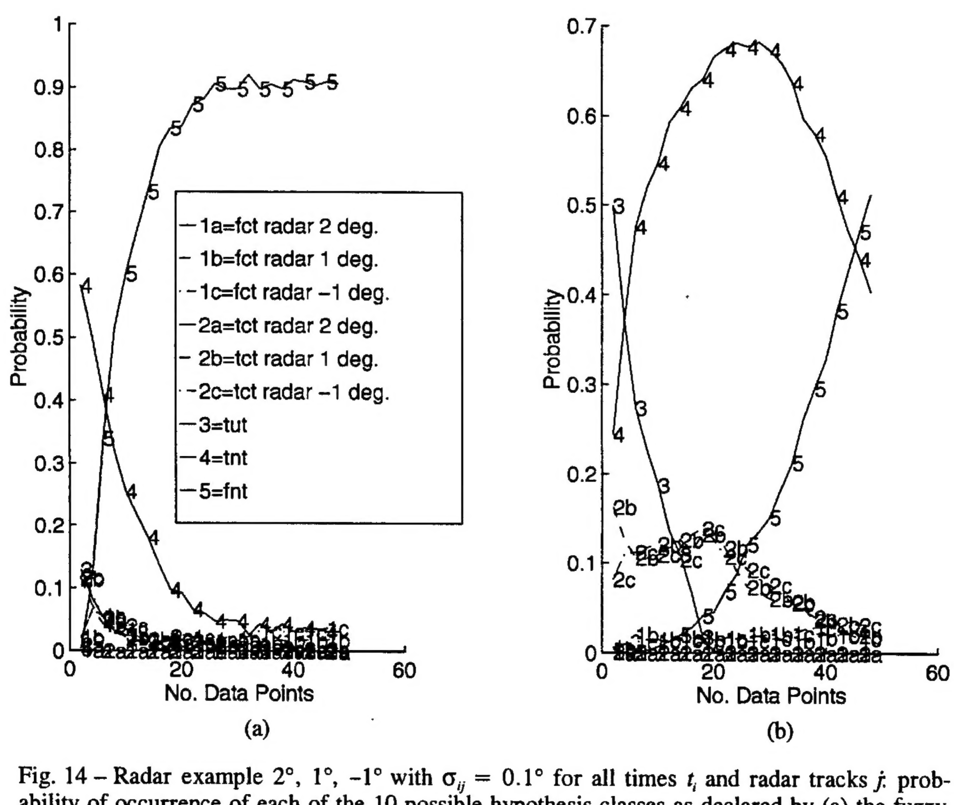


Fig. 14 – Radar example 2°, 1°, -1° with σ_{ij} = 0.1° for all times t_i and radar tracks j: probability of occurrence of each of the 10 possible hypothesis classes as declared by (a) the fuzzy association algorithm and (b) the TW association algorithm

7. FUTURE DEVELOPMENT

The algorithm will be used to examine the effect of variations in data window size. Multiple target cases will be considered. The algorithm will be converted to the computer language C. Finally, the algorithm will become part of a larger ESM tracker currently under development at NRL.

8. CONCLUSIONS

A fuzzy logic algorithm for clustering and associating data measured on different sensors has been developed. It can estimate parameters such as bearing with only a small error. Two subcomponents of the algorithm—known as *superclustering* and *probability augmented superclustering*—can determine the number of targets present in the data. The probability that probability-augmented superclustering determines the number of targets present in the data correctly approaches 100% with as few as 20 data points. Finally, the superclustering procedure can also be used to suppress outliers when dealing with very noisy ESM data.

The fuzzy algorithm's abilities as an association algorithm have been compared to the Trunk-Wilson (TW) association algorithm, a Bayesian philosophy algorithm. In simulations in which noiseless radar data contained truth, the fuzzy association algorithm establishes a firm correlation with one-third to one-half the data required by the TW algorithm. When the noiseless radar data did not contain truth, the fuzzy algorithm outperformed the TW algorithm with only one-sixth the data. When simulated radar data are used with noise of realistic magnitude, the fuzzy algorithm shows almost no deterioration in its performance. The fuzzy association algorithm's ability to make correct decision with fewer data than the TW algorithm is crucial since ESM data are generally sparse, intermittent, and noisy. Finally, the fuzzy association algorithm should be applicable to many different multisensor problems requiring high-quality decisions even though the data are sparse, intermittent, and noisy.

9. ACKNOWLEDGMENTS

This work was sponsored by the Office of Naval Research. The author gratefully acknowledges useful discussions with Mr. Ed Khoury and Dr. Joseph Lawrence III. The author also thanks Gay Polk, Sarah Jones, Kristen O'Neal, Ruth Smith, and Leroy Crispell for administrative help.

10. REFERENCES

1. G.V. Trunk and J.D. Wilson, "Association of DF Bearing Measurements with Radar Track," IEEE Trans AES, **AES-23** (1987).

2. H.J. Zimmerman, *Fuzzy Set Theory and its Applications* (Kluwer Academic Publishers Group, Boston, 1991), p. 11.

3. H. Spath, *Cluster Analysis Algorithms for Data Reduction and Classification of Objects* (John Wiley and Sons, New York, 1980).

4. James F. Smith III, "A Fuzzy Clustering and Superclustering Scheme for Extracting Structure from Data," NRL/FR/5740--96-9844, Dec. 1996.

5. J.C. Bezdek and J.C.Dunn, "Optimal Fuzzy Partitions: A Heuristic for Estimating the Parameters in a Mixture of Normal Distributions," IEEE Trans. Comp. **C-24**, 835-838 (1975).

6. J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms* (Plenum Press, New York, 1981).

7. G. Trunk and J.O. Coleman, "Radar-ESM Correction Decision Procedure," NRL Report 8476, May 1981.

8. E. Khoury, "Advanced ESM Angle Tracker: Vol. I—Tracker Development and Algorithms," NRL/MR/5740-97-9867, to be published.

9. E. Khoury, "Advanced ESM Angle Tracker: Vol. II—User's Guide," NRL/MR/5740-97-7989, Sept. 1997.

10. E. Khoury, M. Hock, B. Weber, C. Fletcher, and B. Cherdak, "Advanced ESM Angle Tracker: Vol. III—Developer's Guide," NRL/MR/5740-97-7990, Sept. 1997.